Journal of Cleaner Production 172 (2018) 1607–1619

Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro

Segmentation analysis of residential water-electricity demand for customized demand-side management programs

A. Cominola [a, *], E.S. Spang [b, c], M. Giuliani [a], A. Castelletti [a, e], J.R. Lund [d], F.J. Loge [b, d]

[a] Department of Electronics, Information, and Bioengineering, Politecnico di Milano, Milan, Italy
[b] Center for Water-Energy Efficiency, University of California, Davis, CA, USA
[c] Food Science and Technology Department, University of California, Davis, CA, USA
[d] Department of Civil and Environmental Engineering, University of California, Davis, CA, USA
[e] Institute of Environmental Engineering, ETH, Zurich, Switzerland

ARTICLE INFO

Article history:
Received 13 March 2017
Received in revised form
15 September 2017
Accepted 18 October 2017
Available online 21 October 2017

Keywords:
Water-energy nexus
Demand management
Customer segmentation
Smart metering
User profiling

ABSTRACT

With increasing water and energy use in the residential sector, due to population growth, urbanization, and climate change, demand-side management (DSM) is essential to complement supply-side interventions to meet future demands and reduce costs. This paper explores how customer segmentation analysis can support customized water and electricity DSM. We contribute a three-phase customer segmentation analysis of over 1000 residential accounts in the Los Angeles County (Southern California) to explore the heterogeneity of residential water-electricity demand profiles and provide insights for coordinated water-energy DSM. Results show that, on the one hand, daily water and electricity consumption are correlated, thus groups of high consumers can be targeted with coordinated water-electricity DSM interventions. On the other hand, the absence of a relevant causal nexus between water and electricity daily load shapes suggests that DSM actions for water should be differentiated from those for electricity. Finally, both objective (e.g., presence of swimming pool) and subjective psychographic features (e.g., conservation attitude) are found to be relevant potential drivers of water-electricity demands. Based on these findings, we propose recommendations for designing a portfolio of mixed customized water-electricity DSM interventions to foster conservation or peak shifting objectives.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Residential water and energy use is expected to increase and change in spatial distribution as a result of population growth (Gerland and Raftery, 2014), urbanization (UNDESA, 2010) and climate change (Cosgrove and Cosgrove, 2012). Considering the increasing marginal cost of building and operating water and energy supply infrastructure, as well as other geographical and financial constraints to infrastructure expansion (Strbac, 2008; McDonald et al., 2014), demand-side management (DSM) strategies are essential to complement supply-side interventions to meet future demands, as well as reduce short and long term costs (Beal et al., 2016).

The effectiveness of DSM options, such as retrofitting incentives (Geller et al., 2006), dynamic hourly pricing (Gaiser and Stroeve, 2014), personalized end-use feedbacks (Fischer, 2008; Carrie Armel et al., 2013), and awareness campaigns (Vassileva and Campillo, 2014), has long been demonstrated in the energy sector. More recently, a growing number of DSM studies for water are exploring the suitability of financial, legislative, educational, and maintenance-based DSM actions (Cominola et al., 2015b) as cost-effective tools to reduce operational costs (e.g., through peak consumption reduction), advance conservation efforts (Mitchell et al., 2013), and improve overall utility performance (Gleick et al., 2003). Interest in these approaches is especially acute in areas suffering from prolonged periods of drought, such as California (Cahill and Lund, 2012). Meanwhile, the literature on water-related energy consumption (Kenway et al., 2013, 2016; Spang and Loge, 2015; Escriva-Bou et al., 2015a) and its implications for linked opportunities to achieve resources conservation and efficiency (Abdallah and Rosenberg, 2014; Escriva-Bou et al., 2015b) is pushing for coordinated water-energy demand management (Cutter et al., 2014).

Regardless of the application, the design of effective DSM

* Corresponding author.
E-mail address: andrea.cominola@polimi.it (A. Cominola).

https://doi.org/10.1016/j.jclepro.2017.10.203
0959-6526/© 2017 Elsevier Ltd. All rights reserved.

strategies relies on a clear understanding of consumer behavior (Jorgensen et al., 2013). In the water-electricity sectors, this has major policy implications. For instance, assessing demand intensity can help target specific groups of users for conservation and customized feedbacks, understanding the time-of-use (Cardell-Oliver et al., 2016) can support the design of dynamic pricing schemes to control peak demand, and identifying consumption determinants (Desmedt et al., 2009; Willis et al., 2013) allows customizing DSM to specifically act on them. The increased deployment of advanced metering infrastructure (AMI, or "smart meters") in both the water (Boyle et al., 2013) and energy (Neenan and Hemphill, 2008) sectors significantly increases the resolution (often hourly) of usage data (Cardell-Oliver, 2013). Data with finer resolution, in turn, support an understanding of use patterns and the detection of anomalies (Britton et al., 2013) in quasi real-time.

As the increasing number of state-of-the-art studies demonstrates (see Cominola et al. (2015b) and references therein), the availability of smart metered data (Gurung et al., 2015) promotes the development of several predictive and descriptive models of residential water or energy consumption to estimate expected future demands based on natural (e.g., temperature and seasonality) and socio-psychographic conditions (i.e., user demography, values, attitudes, interests, or lifestyles, age, number of house occupants, income level, conservation attitude, etc.), and to analyze observed consumption patterns and historical trends. Among descriptive models, customer segmentation and clustering techniques can now discriminate among different daily demand consumption patterns (Espinoza et al., 2005; Verdú et al., 2006), especially in the energy sector. Such techniques allow identifying and characterizing heterogeneous recurring (Kwac et al., 2014) or irregular behaviors (Nizar and Dong, 2009). Recently, some studies extended customer segmentation analysis to the water sector, to identify typical routines in consumption and potential links with household characteristics (e.g., Cardell-Oliver et al., 2016; Cominola et al., 2016).

However, providing accurate representation and complete coverage of the consumption routines of users remains a significant challenge. This paper introduces a customer segmentation analysis to discriminate heterogeneous daily water-energy consumption profiles while filtering out the high levels of noise and variance of smart-metered big data associated with irregular and scattered users' behavior. To our knowledge, no studies currently exist that specifically segment customer types by coupled hourly water and electricity data to inform coordinated water-energy demand management.

We perform a customer segmentation analysis of over 1000 residential accounts in the City of Burbank (in Los Angeles County, Southern California) to identify typical residential water and electricity demand patterns, cluster the accounts with similar water-electricity demand patterns, and provide insights for coordinated water-energy DSM programs. Each account is described by coupled household water and electricity hourly consumption data, as well as psychographic features, i.e., occupants' demography, households characteristics, and water-energy-related attitudes and stated preferences.

Our novel customer segmentation analysis consists of a descriptive modeling method based on hourly water and electricity data. It combines three phases, i.e., (i) eigenbehavior extraction (Eagle and Pentland, 2009) to identify typical water-electricity use routines, (ii) clustering for customer segmentation based on routine similarities, and (iii) factor mapping (Friedman and Fisher, 1999) to infer the potential determinants of targeted water and energy demand profiles. This method enables capturing the heterogeneity of water and electricity consumption profiles among Burbank's accounts, highlighting differences between daily time-of-use of water and electricity, and the characterization of accounts with similar profiles based on psychographic and behavioral features.

From a methodological perspective, the goal of this study is to advance state-of-the-art data mining models of water and electricity user behavior to support the development of customized DSM programs, especially considering the increasing availability of sub-daily sampling resolution data gathered via AMI. Considering the narrative of this study in terms of policy implications, we identify groups of users that can be targeted with conservation actions and other measures (e.g., dynamic hourly pricing schemes) to shift their demand peak. They are those characterized by water and electricity consumption patterns significantly contributing to the total and peak water-energy consumption overall the community of households monitored. Finally, we investigate how a portfolio of diverse water-electricity DSM strategies can be customized to act on objective and/or personal and subjective features, which represent potential demand determinants for the targeted groups of users.

2. Material and methods

2.1. Case study of Burbank, California

Our case study focuses on the City of Burbank, an urban area of approximately 103,000 residents (US Census Bureau, 2010) in Los Angeles County in Southern California (Fig. 1). Burbank is in the South Coast hydrologic zone of California, which has a Mediterranean climate with distinct seasonal precipitation patterns of winter rains (75% of precipitation between December and March) and dry, hot summers. While the South Coast is only one of ten hydrologic zones across the State, nearly 53% of the State's population lives in the region and average water use in 2010 was approximately 188 gallons per capita per day, i.e., over 700 L per capita per day (California Department of Water Resources (DWR), 2014).

The area at the time of data collection for this study (2015) was facing a period of prolonged drought condition. Moreover, it is worth mentioning that outdoor water usage contributes on average to more than half of household water use in South California, and that water-related electricity use for domestic water heating is generally low, equal to approximately 2% of total household electricity use in California (see Fig. 2 for an overview of average residential water and electricity share by appliance type).

Burbank Water and Power (BWP), a municipal utility, serves the City of Burbank as joint water and electricity provider. BWP was uniquely well-suited as a case study given their deployment of AMI systems to collect both water and electricity consumption data from their service territory area at hourly intervals. Further, as a combined water and electricity utility, they have a unique understanding of the potential for shaping both water and energy DSM strategies and a depth of experience implementing water and energy rebate programs and incentives.[1]

In this study, we perform user profiling and customer segmentation on an anonymous, processed dataset of hourly water and electricity consumption data gathered for 1107 residential accounts in Burbank during the period June, 28th - December 8th, 2015, with each account having 4584 hourly readings (191 days), and the full dataset having 5,074,488 data points. Our sample represents about 6% of the roughly 18,000 total single-family residential accounts serviced by BWP. The 1107 accounts were selected from the broader residential population for this study based on

[1] More information on the conservation programs here: https://www.burbankwaterandpower.com/rebates-and-incentives.

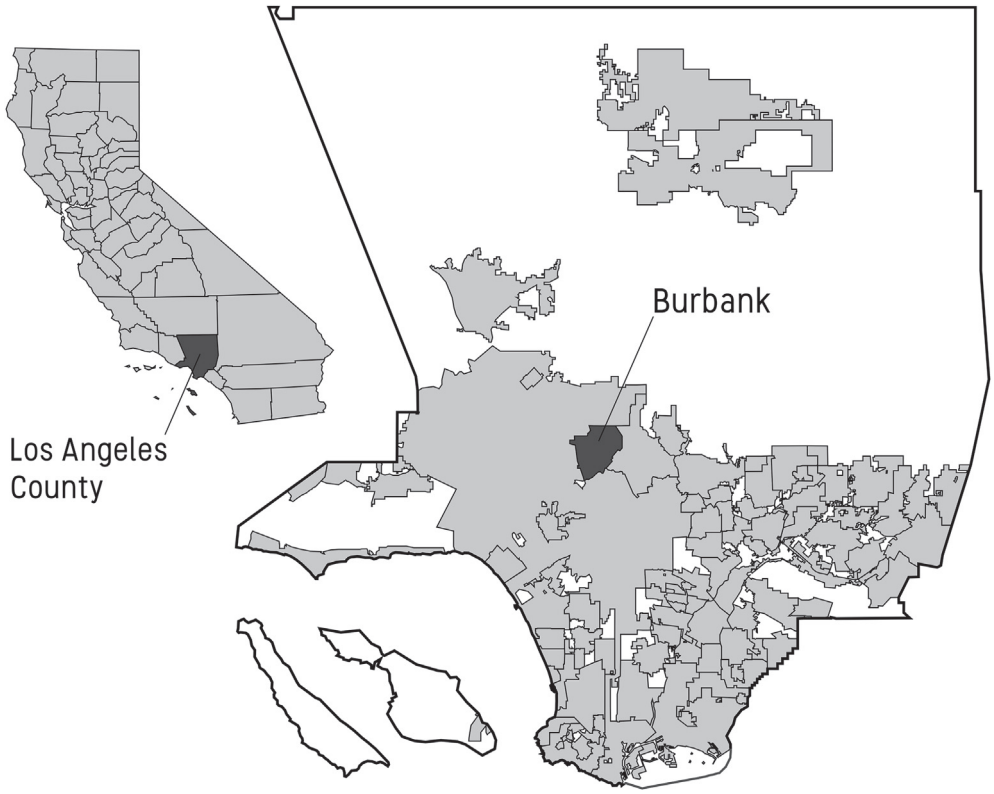


Fig. 1. Los Angeles County (California) with highlight of the Burbank Area.
Figure adapted from: https://commons.wikimedia.org/wiki/File:LA_County_Incorporated_Areas_Burbank_highlighted.svg.

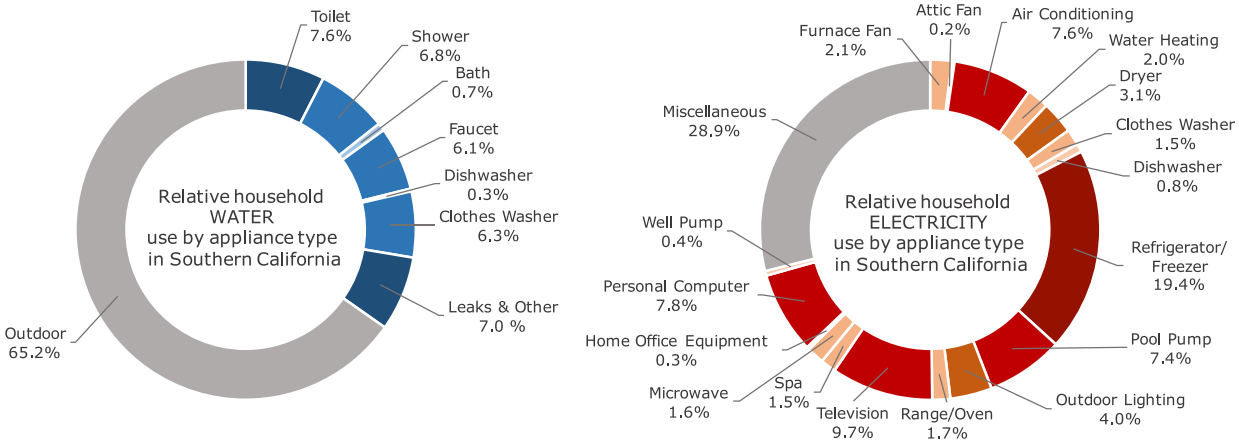


Fig. 2. Relative household water use by appliance (donut chart on the left) and relative household electricity use by appliance type (donut chart on the right) in Southern California. Water use percentages by appliance type have been evaluated as an average of the water use data per utility and end-use across the utilities located in South California (Las Virgenes MWD, Los Angeles DWP, IRWD, San Diego City, San Diego County) reported in Escriva-Bou et al. (2015a). Electricity percentages have been obtained from the average annual electricity consumption by appliance for single family homes in California reported in Palmgren et al. (2010).

their participation in an entry survey linked to a behavior-based water conservation initiative implemented by BWP in early 2015. The opt-in, web-based survey provided information on several key psychographic features from participating users, including: occupant demography, households characteristics, and water/energy consumption attitudes and stated preferences. We considered 50 psychographic features included in the survey (WaterSmart (2015), see Appendix for the complete list).

2.2. Customer segmentation method

We developed a three-phase method (see Fig. 3) to identify representative (i.e., recurring) water and energy user profiles, to perform consumer segmentation based on the profiles, and to infer the potential determinants of targeted water and energy consumption profiles.

Each phase is implemented as follows:

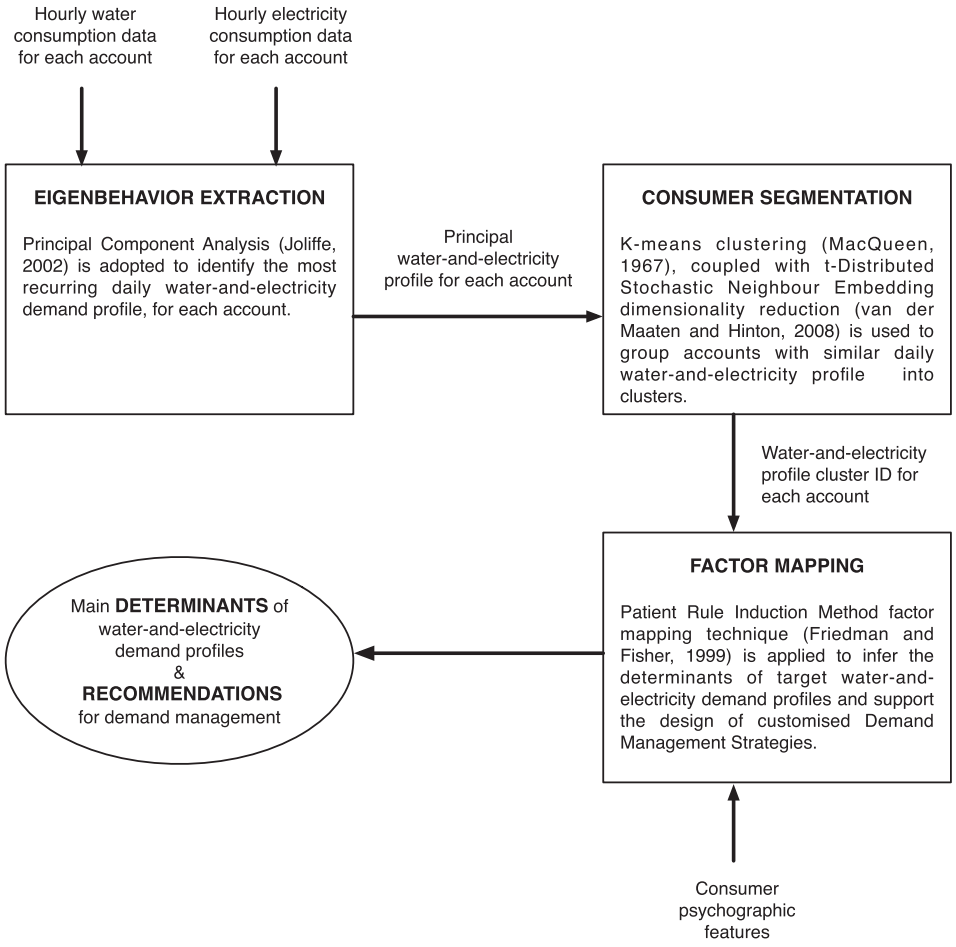

Hourly water consumption data for each account
Hourly electricity consumption data for each account
EIGENBEHAVIOR EXTRACTION
Principal Component Analysis (Joliffe, 2002) is adopted to identify the most recurring daily water-and-electricity demand profile, for each account.
Principal water-and-electricity profile for each account
CONSUMER SEGMENTATION
K-means clustering (MacQueen, 1967), coupled with t-Distributed Stochastic Neighbour Embedding dimensionality reduction (van der Maaten and Hinton, 2008) is used to group accounts with similar daily water-and-electricity profile into clusters.
Water-and-electricity profile cluster ID for each account
FACTOR MAPPING
Patient Rule Induction Method factor mapping technique (Friedman and Fisher, 1999) is applied to infer the determinants of target water-and-electricity demand profiles and support the design of customised Demand Management Strategies.
Main DETERMINANTS of water-and-electricity demand profiles & RECOMMENDATIONS for demand management
Consumer psychographic features

Fig. 3. Flowchart of the three-phase customer segmentation method.

i) *Eigenbehavior extraction.* The first step aims at characterizing each account by its most recurring coupled daily water and electricity demand profile. Inputs for this analysis are hourly water and electricity consumption time series. Our approach builds on the idea proposed by Eagle and Pentland (2009) to identify the time- and location-based routines of 100 individuals monitored via smartphone communications. Here, we extracted recurring coupled water and electricity daily demand profiles from the hourly water and electricity datasets by non-centered Principal Component Analysis (PCA) (Jolliffe, 2002), i.e., variables are not scaled by subtracting their mean before computing PCA because it is performed on binary data as in Cominola et al. (2016).

PCA is a dimensionality reduction technique which searches for linear combinations of the original variables (i.e., water and electricity consumption for each hour of the day) such that the coefficients of the output combinations (the principal vectors) form a low-dimensional sub-space defined by vectors that explain the maximal variance in the original data. The extraction of principal components from behavioral data defines a set of vectors spanning the "behavioral space" of monitored individuals, characterizing their behavioral variation in time. These components, called eigenbehaviors, are computed as the eigenvectors of the covariance matrix of behavior data, where the vectors with high weights represent a type of recurrent behavior, i.e., a routine. Usually, a few principal components can explain much of the variance of the original variables, enabling dimension reduction. In this work, the first eigenbehavior accounts for most behavioral data variance in user's consumption, so we restrict our analysis to the first eigenbehavior of each account.

For our case study, we labeled the water data prior to eigenbehavior extraction using the following logic: *Zero demand* includes values of water consumption equal to 0 cubic meters per hour, which are over 60% of all Burbank's accounts hourly water readings; *Low-medium demand* includes values in the range $(0, 47.29]$ L/h, where 47.29 L/h represents the amount of water used for a 5-min shower with a flow of 9.46 L per minute (i.e., 2.5 gallons per minute) (DeOreo et al., 2016); *Medium-high demand* includes consumption values higher than 47.29 L/h, typical, for instance, of long showers, some clothes washer programs, and outdoor uses. In contrast to the intermittent water usage demand, hourly electricity consumption data are always positive, i.e., some electricity is always being consumed by the household. We labeled electricity usage data as follows: *Low demand* includes values of electricity consumption lower or equal to 0.36 kWh, i.e., the 25-th percentile of all Burbank account hourly electricity readings, and is assumed to include low-level consumption for plug loads, stand-by appliances, or lights; *Medium demand* includes values in the range $(0.36, 1.5]$ kWh, with the upper bound of 1.5 kWh representing the 75-th percentile of the observed hourly electricity readings, and comprising electricity uses for larger appliances like dishwashers, microwaves, and toaster ovens; *High demand* includes

consumption values exceeding 1.5 kWh representing the highest electricity loads, e.g., air conditioners, space heating, clothes dryers, and electric ovens.[2] In practical terms, coupling this data labeling with eigenbehavior extraction means that cumbersome hourly smart-metered time series composed of thousands of data points can be represented with one relevant daily consumption pattern, composed of only 24 data points (one for each hour of the day) for each water/electricity labeled category, which represents the most relevant consumption routine of each user.

ii) *Consumer segmentation.* After data dimensionality is reduced and the principal behaviors (i.e., eigenbehaviors) are estimated via PCA for each account, we cluster the accounts based on the similarities of their first eigenbehavior. This step produces an understanding of the core "types" of daily water and electricity consumption routines detected across the sampled community. This is relevant for assessing the similarities in the main daily consumption routines across the users, quantifying the relative proportion of each cluster's water and electricity demand to total usage, and targeting clusters with customized demand management strategies. The method used for this classification was a sequential implementation of t-Distributed Stochastic Neighbor Embedding (t-SNE) (van der Maaten and Hinton, 2008; van der Maaten, 2014) followed by traditional K-means clustering (MacQueen, 1967). The t-SNE method allows for dimensionality reduction of high-dimensional objects into two-dimensional points. For our case, we applied it to the first eigenbehavior extracted to further simplify dimensionality (down to two data points for each eigenbehavior) and enhance data separability into clusters. We then applied the K-means clustering analysis to compute the final clusters, i.e., automatically group the data in the *k* clusters of similar water-electricity usage behaviors.

iii) *Factor mapping.* In the last phase we infer the most likely determinants of water-electricity eigenbehaviors of targeted accounts clusters (e.g., high consumers) from a list of psychographic features. Identifying water and electricity demand drivers is key for customizing DSM strategies that can effectively shape consumer response. The input list of candidate drivers included consumers' psychographic features related to occupant demography, household characteristics, and water- and energy-related attitudes and stated preferences. We extracted the most likely demand drivers from the candidate list by mapping the latter for targeted accounts' clusters via the Patient Rule Induction Method (PRIM) (Friedman and Fisher, 1999). PRIM is a factor mapping technique used recently for scenario discovery under uncertainty (Kwakkel and Jaxa-Rozen, 2016) as it enables identification of ranges and combinations of uncertain input variables that result in similar ranges of output variables (Kwakkel and Pruyt, 2015), suggesting that the selected inputs are likely to be correlated with the targeted output range.

The resulting combination of eigenbehavior extraction, customer segmentation, and factor mapping allows proposing recommendations for the design of customized and coordinated DSM strategies for water-electricity conservation or peak shifting.

3. Results

3.1. Analysis of coupled daily water-electricity consumption profiles

This section analyzes coupled water-electricity daily consumption profiles for the BWP study population accounts. Eigenbehavior analysis and clustering lead to consumer segmentation according to steps (i) and (ii) of the method described above (Fig. 3).

The first water-electricity eigenbehavior for every account is visualized as an individual matrix column in Fig. 4. Each column represents the eigenbehavior coefficients, subdivided into three labeled levels of water consumption (*Zero*, *Low-medium*, and *Medium-high*) and three labeled levels of electricity consumption (*Low*, *Medium*, and *High*, see Section 2.2 for details on label definition). Each consumption level has 24 coefficients, one for each hour of the day. Color intensity is proportional to coefficient magnitude: darker colors have higher weight given for a level of consumption for the given hour.

Fig. 4 shows that a variety of eigenbehaviors exists across the Burbank households in the study. We identified 18 profile clusters, shown as P1, P2, …, P18 in the figure. Among the profiles we can identify typical behaviors which differ in eigenbehavior coefficients. For instance, P1 has both high to medium-high water and electricity consumption. The increased water and electricity demand of these users could be driven by large house size, multiple occupants, and/or intensive end-uses like swimming pools, intensive landscaping, and large-scale air conditioning. On the other extreme, P17 includes very low consumers, including perhaps single-occupant small houses/vacation houses. Profiles such as P14 are in the middle, with the highest coefficients assigned to the daytime hours for medium levels of water and electricity use.

Beyond the scale of usage, profiles that differ by the shape of their load curves can also be identified. For example, while most profiles show peaks in water use during early morning and evening hours, and high electricity consumption in the afternoon, profiles like P4 and P18 show overall low water and high electricity use concentrated in working hours. This could be due to some home business, even though we lack data to verify this hypothesis. Overall, results show that most consumers do not use water at night (dark color coefficients show *Zero water consumption* between midnight and 6 a.m.) and much of the day, confirming the 60% of zero hourly water consumption in the raw data. Generally, there is no water use when people sleep or are not at home (e.g., during working hours). Conversely, about 50% of the profiles (P1 to P9) show high coefficients for medium and high electricity consumption both day and night. This high base load in working and night hours might be due air conditioning, programmable appliances (such as clothes dryers) or many appliances always plugged in with *on* or *stand-by* mode.

Given the mismatch between water and electricity eigenbehaviors in terms of time-of-day usage and consumption peaks, we conclude there are not clear cause-effect mechanisms between residential water and electricity end-uses in our case. Indeed, this is re-enforced by the fact that water is heated by gas in most of Burbank's accounts.[3] Consequently, water consumption is not the direct driver of electricity consumption, and vice versa. Thus, we conclude that both demands are likely to be driven by other features of the houses and their occupants, including behavioral features and attitudes.

We provide an alternative visualization of the different water-electricity eigenbehaviors in Fig. 5, where the median

[2] Source for examples of appliance classification in the three intervals: http://www.siliconvalleypower.com/for-residents/save-energy/appliance-energy-use-chart.

[3] From survey data, around 91% of the accounts considered in this study declared they use gas heating.

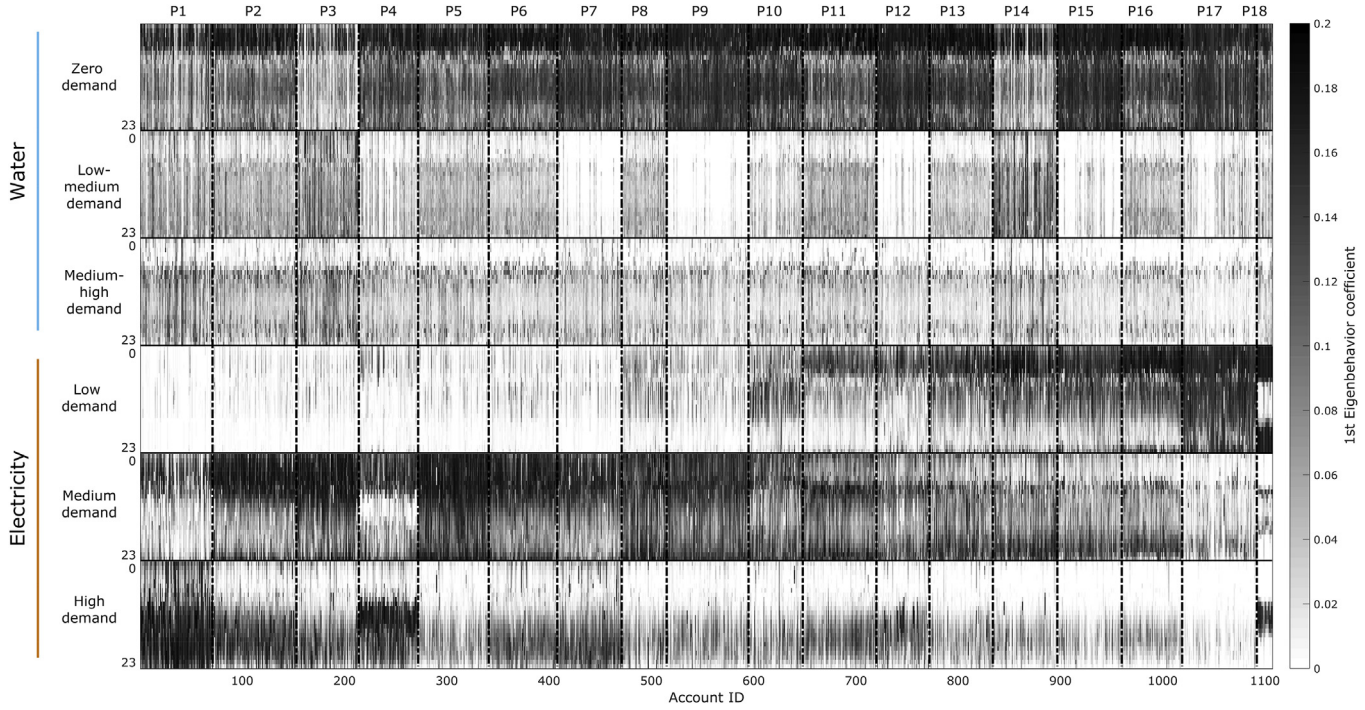


Fig. 4. Water-and-energy eigenbehavior coefficients and clusters. Each column matrix represents coefficients of the first eigenbehavior for one account (each account is identified by means of an account ID, see x-axis). Eigenbehavior coefficients are reported for diverse labeled levels of water and electricity consumption over 24 h (see y-axis). Color intensity is proportional to coefficient magnitude, as represented in the colorbar on the right side. Eigenbehaviors are clustered in 18 groups based on their similarity: clusters are identified by the labels P1, P2, …, P18 on top of the figure and the vertical dashed lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

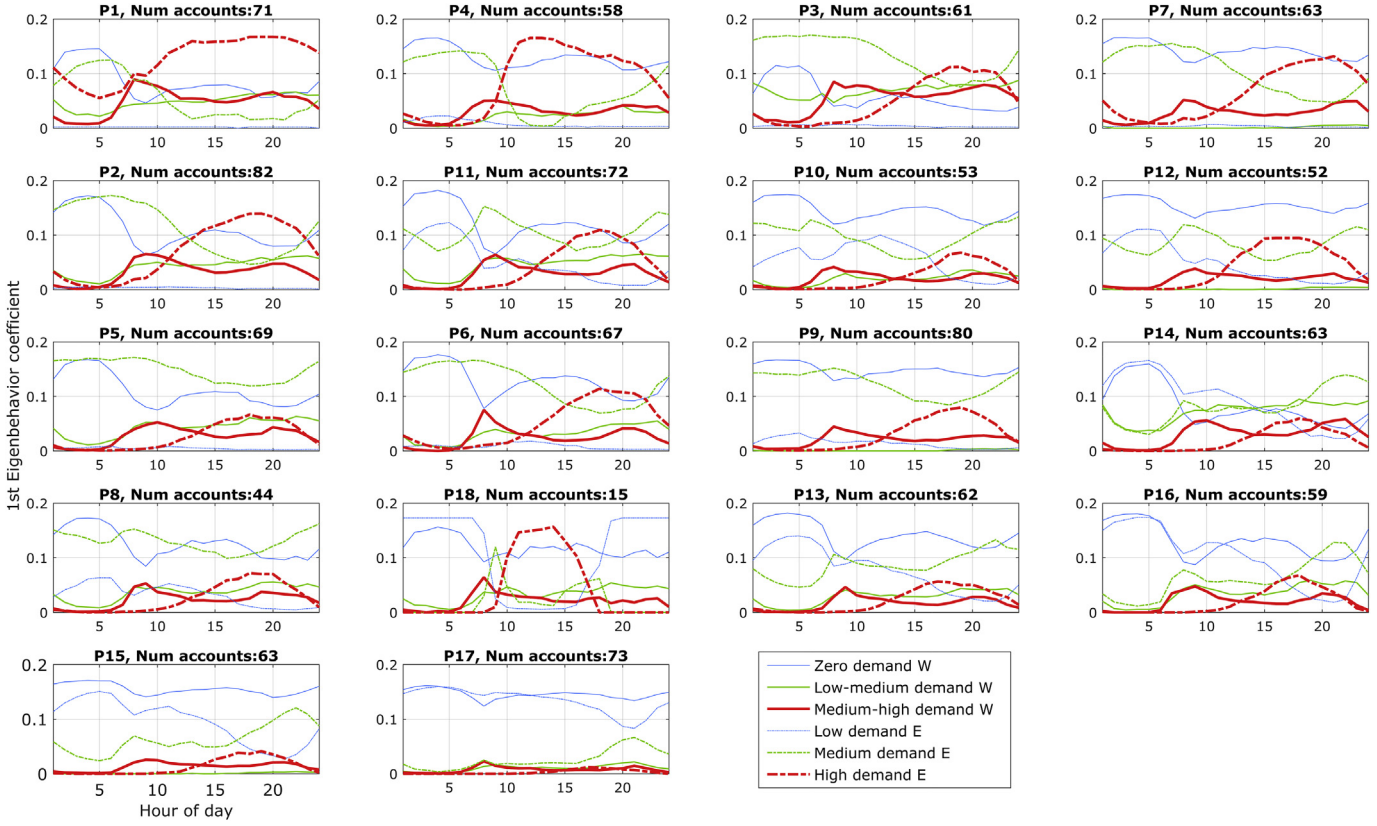


Fig. 5. Median weights of the first eigenbehavior for each profile P1, P2, …, P18. Profiles are sorted in decreasing order of total per-capita water consumption during the observation period (from top-left to bottom-right). Solid lines refer to water, dashed lines to electricity. Colors are referred to different consumption labels. On top of each subplot the number of accounts for each profile ("Num accounts") is reported. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

eigenbehavior weights of each profile P1, P2, …, P18 are represented with solid (dashed) colored lines for the 3 water (electricity) consumption levels defined above, and profiles are sorted in decreasing order of total water consumption during the observation period. Also, the number of accounts for each profile is reported. Median profiles confirm the overall mismatch between water and electricity time-of-use and consumption peaks, i.e., the solid red line indicating high water consumption shows two peaks, the highest in early morning hours (6–10 a.m.) and a second in the evening (6–11 p.m.), while the dashed red line, indicating high electricity consumption shows a bigger peak that usually starts in late morning/early afternoon hours and then lasts till late evening hours. Finally, behavioral differences emerge among high-consuming water profiles, denoting two kinds of water-intensive behaviors. Profiles such as P1 and P3 describe users that regularly use high amounts of water during day hours, demonstrated by the larger values of the solid red (high usage) and green (medium usage) lines relative to the solid blue line (low usage) during peak hours. In contrast, profiles like P2 and P4 describe the behavior of users that only occasionally use high amounts of water during day hours (evidenced by the larger values of the solid blue line relative to the solid green and red lines for all hours), but total water consumption during the monitored period comparable to that of P1 and P3.

3.2. Recommendations for the design of customized DSM strategies

The characterizations of profiles we detailed in the previous section in terms of time-of-use of water and electricity and frequency of high demands are essential to support the design of DSM strategies based on pricing. Indeed, increasing block tariffs (i.e., rates which increase unit price as more water is used (Spang and Loge, 2015)), penalizing high water and electricity use may be effective at incentivizing conservation by consumers having profiles of "flat" high consumption like P4. Conversely, dynamic hourly pricing schemes may be better for reducing the "peakiness" of consumption (and thus reducing utilities costs during peak hours (Beal et al., 2016)) as in profile P3. In practice, offering selected pricing schemes to different groups of customers is likely unfeasible, given the inherent inequality of this approach as well as the more general social and political challenges of adjusting water rates (Spang et al., 2015). Still, utilities can exploit the information of different customer profiles to design a common pricing scheme (e.g., either just block tariffs or hourly dynamic pricing) suitable to target the most critical profiles and maximize overall benefit in terms of conservation and cost savings. Moreover, utilities can adapt the selected pricing scheme over time, based on observed behavioral changes and modifications of consumption intensity and timing.

The main heterogeneous typologies of behaviors discussed in Fig. 4 can be recognized in the above comments, and are summarized in Table 1 with respect to DSM interventions based on time-of-day water and electricity use.

In the remainder of this section, we cross-analyze consumption profiles and users' psychographic features (see Appendix) through factor mapping analysis to formulate further recommendations for DSM. Given that most candidate features for factor mapping were retrieved as consumer psychographic data collected via a water-related survey to consumers, the results are skewed towards characterizing water demand. Yet, in the next paragraphs we show that our results provide key insights for potentially improving energy DSM as well.

3.2.1. Water-energy conservation programs

Despite the variation in timing of water and electricity consumption routines, average water demand and average electricity demand of each profile show a linear correlation of 0.93 (see Fig. 6). Based on this comparison, we identified the set of profiles with largest water and electricity consumption demand. Such profiles, visualized with a dark and light blue boxes in Fig. 6, should be considered high-priority targets for water and energy conservation actions, as they contribute most to water-electricity demand (profile P1 accounts for 9.64% total water use and 13.56% total electricity use, despite it includes only 6.4% of all users. Profiles P1, P4, P2, P7, and P3 account together for 39.71% total water use and 47.15% total electricity use, while including only 30% of all users).

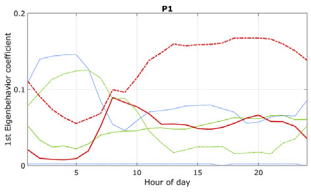
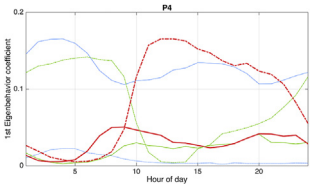
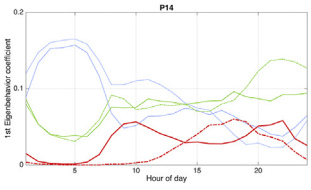
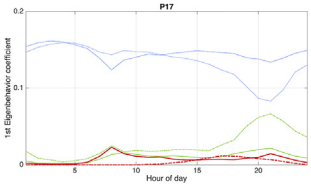
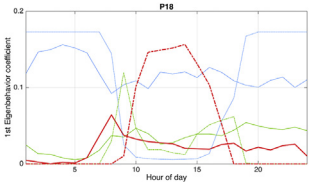
Targeting the top consuming profile (P1), as well as the next top 5 profiles (P1, P4, P3, P7, P2), we run factor mapping analysis to explore potential determinants of their demands. Results are summarized in Table 2, under the *Water conservation* DSM scenario.

Not surprisingly, results show that the selected feature, i.e., the main feature identified as critical for explaining the profile with higher water and energy consumption (P1) is the presence of either a swimming pool, or a hot tub, or both, as highlighted in bold in Table 2 under column *Feature range*. Around 75% of the users belonging to profile P1 declared to own a swimming pool, a hot tub, or both, while the remaining declared to own "Neither" a pool or a hot tub. This finding is in line with previous studies (Mayer and DeOreo, 1999) that demonstrated that not only are swimming pools one of the highest consuming end-uses of water (homes with swimming pools use more than twice as much water outdoors than homes without swimming pools), but also that those households with a swimming pool on average showed a more consuming attitude for other water end-uses, as well as a higher rate of water leakage, up to 55%.

This may also partly explain the high electricity demand of targeted profiles, contributed by pools and hot tubs for pumping and filtration. Even though we do not have precise data on electricity use for pool pumping and filtration across our sample accounts, it is estimated that 7.4% of residential electricity used on average in California is for pool pumps (see Fig. 2). Yet, with the available information, we cannot *a priori* exclude that high electricity demand may be explained by a higher number of house occupants or a more consuming attitude by household occupants. The selection of a specific outdoor end-use as a main potential determinant of demand is key for demand management. Indeed, this outcome opens up diverse options for the design of customized DSM. Feedback to users tailored on conserving on pool/hot tub use and refilling can in principle reduce water (and partially energy) demand. However, this is not the only option for conservation: given that the selected outdoor end-use involves high consumption rates, utilities might employ block tariffs schemes penalizing high consumption habits or season and time-dependent restrictions regulating the use of water for outdoor end-uses during scarcity periods.

Extending the analysis out to the top 5 profiles (P1, P4, P3, P7, P2), the previous result is confirmed and pool/hot tub presence remains an important driving feature of consumption. Additionally, demographic aspects (i.e., number of house occupants) emerge as relevant, as hypothesized in the discussion of the water-energy profiles estimated in the previous section. Water demand also increases with the number of house occupants (Cominola et al., 2015a), meaning that some top-consuming accounts show high demand just because of their size and the overlapping routines of multiple users. In that case, as the total consumption is given by the sum of individual users consumption, the presence of multiple users within the same household can be exploited as an opportunity to reduce consumption waste. For instance, this can be pursued through customized feedbacks aimed at increasing efficiency in water and energy use via economies of scale in the household (e.g.,

Table 1
Characterization of main heterogeneous typologies of behaviors.

Profile ID	Profile Name	Characteristics and implications for DSM
	High-regular water-electricity consumer	- Users regularly use high amounts of water and electricity during day hours. High "peakiness" in peak hours (red solid and dashed lines are higher than blue and green lines during peak hours). - Dynamic hourly water and electricity pricing schemes may reduce demand "peakiness". Increasing block tariffs, efficient devices, and conservation feedbacks can incentivize overall conservation.
	High-occasional water consumer	- Users only occasionally use high amount of water during day hours (green and red solid lines are lower than the blue solid line). Total water demand close to that of *High-regular consumers* and less marked peaks than *High-regular consumers*. High level of electricity use. - Increasing block water tariffs, efficient devices, and conservation feedbacks can incentivize overall water conservation.
	Average water-electricity consumer	- Mostly, medium levels of water (solid green line) and energy use (dashed green line) during daytime hours. Seldom higher peaks during peak hours. - Not high priority targets for DSM interventions. Still, customized feedback and efficient devices can hep reducing water-electricity waste.
	Low consumer	- Low consuming profiles, high coefficients to *Zero* water consumption (blue solid line) and *Low* electricity use (blue dashed line). - Not a priority targets for DSM interventions.
	Daytime consumer	- Low water use and high electricity use concentrated in working hours. Perhaps, users with this profile run a home business. - Given the overall low water and electricity consumption, not priority targets for water and electricity conservation actions. Potential priority targets for electricity peak shifting programs.

running the clothes washer or the dishwasher only when full) or more efficient devices (Beal et al., 2011). As the number of house occupants influences both water and electricity consumption, feedback-based DSM actions are potentially effective to foster an efficient use of both water and electricity.

We repeated the factor mapping analysis on only users without a pool or a hot tub, including users progressively from the top 1 (P1) to the top 5 (P1, P4, P3, P7, P2) profiles, to see whether for any subset of targeted clusters other features are selected as demand determinants. Interestingly, attitudes and subjective perceptions become relevant when considering users belonging to the P1 *w/o pool* and (P1, P4, P3) *w/o pool* clusters. Users belonging to those top consuming clusters declared a medium-to-low sensitivity towards water price and a medium-to-low environmental interest towards saving water. While taking into account that there is often a difference between what consumers declare (stated preference) and how they actually behave (revealed preferences) (Beal et al., 2013), this information is essential to DSM, as programs based on pricing will likely be ineffective at incentivizing conservation. Further, the effectiveness of water-energy awareness and education efforts will depend greatly on the flexibility of these users' perspective on the importance of conservation.

3.2.2. Water peak shifting programs

The objective of DSM is not only to reduce water and energy demand, but also to reduce utility costs during peak hours (Beal et al., 2016; Gurung et al., 2016). As such, we explored a scenario focused on reducing utility costs due to morning peak water use by repeating the factor mapping process to target those profiles contributing to most of the water demand between 6 and 10 a.m.

While results mainly confirm the findings discussed above for conservation programs, two interesting results emerge when considering progressively up to the top 4 profiles targeted for peak shifting, i.e., P1, P3, P4, and P2 (see the last line in Table 2). First, the influence of a swimming pool is reinforced. However, there is more nuance to the conclusion as the critical range of that feature is now restricted to having "Both a pool and a hot tub" and "Pool only", because the peak flow of a swimming pool (or its combination with the peak flow of a hot tub) is larger than that of only a hot tub. Secondly, the relevance of personal attitudes is enhanced, as

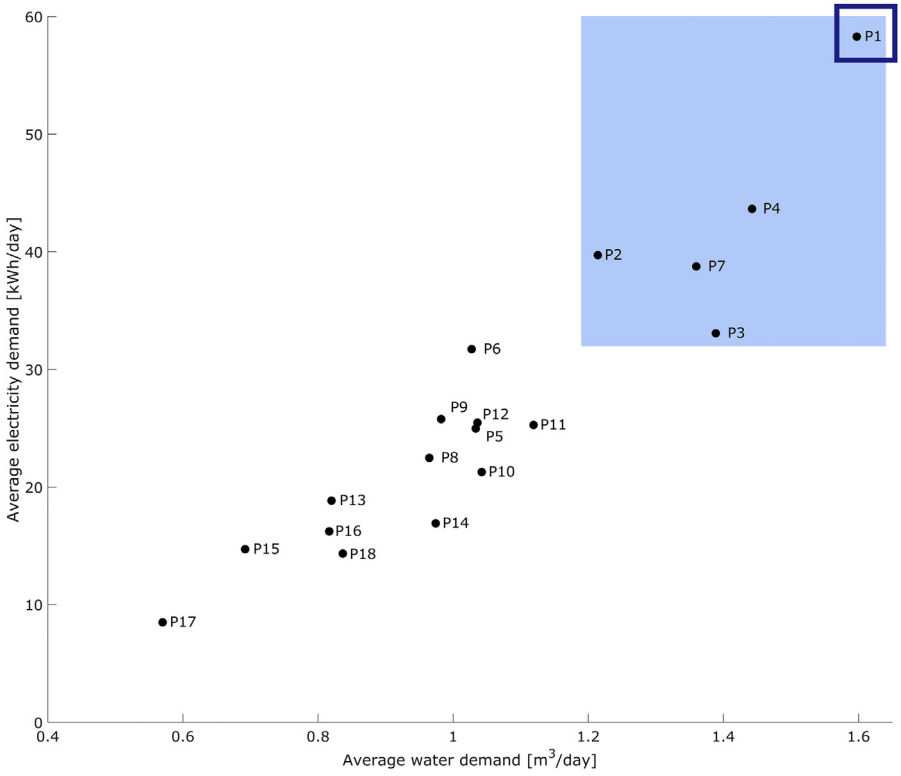


Fig. 6. Average per-household daily water and electricity consumption for each consumer cluster. Each scatter point represents a cluster of consumers. Clusters falling inside the dark and light blue boxes are potential targets for water and energy conservation interventions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

targeted users declared a medium-to-low trust in believing their water utility helps them understand their water use. This is an important indication for DSM and utilities. Indeed, this suggests that DSM should include education-based actions. Further, on the utility side, services that provide consumers' with accessible and detailed information on their consumption, as well as incentives to

Table 2
Factor mapping results for customers' clusters to target for water demand management programs. Bold font in the last column highlights the critical range for the selected feature reported in the third column.

DSM scenario	Target profiles	Selected feature	Feature range
Water conservation	P1	Pool/hot tub presence	**Both a pool and a hot tub** **Pool only** **Hot tub only** Neither
Water conservation	P1, P4, P3, P7, P2	Pool/hot tub presence	**Both a pool and a hot tub** **Pool only** **Hot tub only** Neither
		Number of house occupants	3–8
Water conservation	P1 w/o pool	Interest in spending less on water bill	**Strongly disagree** **Somewhat disagree** **Neither agree or disagree** **Somewhat agree** Strongly agree
Water conservation	P1, P4, P3 w/o pool	Saving water helps the environment	**Strongly disagree** **Somewhat disagree** **Neither agree or disagree** **Somewhat agree** Strongly agree
Peak shifting	P1,P3,P4,P2	Pool/hot tub presence	**Both a pool and a hot tub** **Pool only** Hot tub only Neither
		Number of house occupants	3–8
		User agrees the utility makes it easier for him/her to understand his/her water use	**Strongly disagree** **Somewhat disagree** **Neither agree or disagree** **Somewhat agree** Strongly agree

promote efficient behaviors (Novak et al., 2016), can increase two-direction communication between utilities and users. This, supported by detailed data from smart meters, can increase transparency and data sharing. These latter are key to allow users to better understand their demands, as well as the water/electricity usage they pay for, and ultimately increase their trust toward the suppliers.

4. Discussion

Overall, our findings suggest that the proposed method for water-electricity consumer segmentation is suitable for identifying clusters of typical residential water and electricity demand patterns from data collected via AMI. Using PCA to extract recurrent behaviors from large high-resolution datasets has a number of advantages relative to the simple computation of data averages along each dimension. PCA allows for extraction of multiple primary and secondary behaviors, each explaining an amount of data variance. The amount of variance explained by each component allows a quantitative understanding of its relevance, and thus, the identification and selection of primary behaviors. Moreover, as components are evaluated along directions of maximal explained variance, they are guaranteed to accurately represent data heterogeneity, while averages can be strongly biased by recurring extreme values. In our results, the first eigenbehavior explained on average more than 60% of the total data variance, thus we considered only one primary behavior for our analysis.

The sequential application of t-SNE and K-means further enhanced the data dimensionality reduction rationale behind the PCA component of our method. The two-dimensional data representation obtained by t-SNE allowed us to better visualize data separation and identify the proper number k of clusters to set as input to K-means. Finally, the last component of our method, i.e., PRIM has been demonstrated to identify the heterogeneous psychographic features (and their range) that are most likely to describe consumers belonging to some targeted account clusters (e.g., high consumers).

The narrative of our results in terms of demand management demonstrates that our customer segmentation provides key insights to coordinated water-energy DSM programs for resources conservation or demand peak shifting. On the customer side, results show that actual household features and personal attitudes or perceptions contribute to the observed patterns of their water and electricity consumption. Utilities and planners should consider these findings to differentiate their portfolio of water-electricity DSM and tailor according to the characteristics of the user community and the overall DSM objectives.

Yet, as we consider this study as a starting point for further applications of our customer segmentation method, it is worth discussing the potential case study-specific limitations of the analysis. First, our sample population suffers from self-selection bias based on their election to respond to an online survey. Thus, the patterns observed in our sample are not directly transferable to the overall population of residential consumers in the utility territory. Future studies should seek to randomize household selection for improved extrapolation of study results. Moreover, we acknowledge that gathering comprehensive databases of user psychographic information is often limited by data availability and user privacy issues, as well as costs for frequent survey campaigns. Web-based surveys such as the one adopted to gather the psychographic information used in this study, as well as interactive digital tools possibly coupled with incentive mechanisms, may foster users' to share such data and keep them updated, enabling available datasets to grow overtime.

Second, the application of our analytic framework to communities of users monitored for longer than 6 months would allow including the effect of meteorological conditions and seasonality on consumer behavior into the analysis. Having a larger number of entire weekly periods monitored at high resolution would also allow distinguishing differences in water and energy consumption behavior between weekdays and weekends.

Finally, while the application of our customer segmentation can be easily replicated onto contexts other than Burbank, the outcomes in terms of water-electricity use profiles, water-electricity nexus, and demand management implications should be carefully considered as context-specific. For instance, our finding that water and electricity lack a clear cause-effect nexus and time-of-use synchronism across the considered Burbank accounts (mainly because of gas-based water heating) may not stand for other US or global areas with a different energy mix and a greater penetration of electricity-based water heating.

5. Conclusions

In this paper, we contribute a customer segmentation analysis of over 1000 residential accounts in Burbank (Southern California) to explore heterogeneity in typical residential water-electricity demand profiles, interpreting them in terms of users' behavior, and providing insights to coordinated water-energy DSM. Our customer segmentation analysis integrates three phases, i.e., (i) eigenbehavior extraction for the identification of typical consumers profiles from smart metered hourly water-energy time series, (ii) clustering for customer segmentation based on profile similarities, and (iii) factor mapping to infer the potential determinants of targeted profiles to inform the design of customized DSM.

Results demonstrated that our method can capture heterogeneity in water-electricity consumption profiles and routines over a community of users, highlighting differences between daily time-of-use of water and electricity, and allowing for the characterization of accounts based on psychographic and behavioral features. Moreover, factor mapping can discriminate among features most likely to influence water and energy demand for targeted groups of users.

In terms of key implications for DSM, first, we found that daily water consumption profiles differ from electricity profiles, in terms of time-of-day usage and consumption peaks. This is mainly because water is mostly heated by gas in Southern California, so there is not a clear causal nexus between water and energy demand, and vice versa. This means that DSM strategies for changing daily demand patterns (e.g., hourly dynamic tariffs) should be differentiated between the two sources. At the same time, despite their different load shapes, correlation exists between total water and energy consumers' use. This suggests that common groups of users can be targeted for water-energy DSM. This analysis would be strengthened significantly by the availability of consumption data measured or estimated at the end-use level. This would allow for more detailed characterization and differentiation of users' behaviors, as well as a more nuanced assessment of the influence of users' psychographic features on specific end-uses.

By coupling the meter data with survey data we determined that objective features (e.g., presence of a swimming pool), as well as subjective and personal ones (e.g., attitudes toward resources conservation) are relevant for characterizing users by their consumption habits and routines. The type of features emerging from the factor mapping phase is essential for customizing DSM strategies to match each user's consumption determinants. Yet, we should also consider that there may often be a mismatch between declared and observed users' preferences, attitudes, and behaviors. This can affect the effectiveness of demand management strategies, so repeated data analysis and behavioral change monitoring are

needed to adapt such strategies depending on the observed impacts over time. Ideally, a joint behavioral experimental study comprising both a data analysis phase and a practical phase with the actual implementation of DSM strategies on controlled groups of accounts, designed on the basis of the output provided by our method, would be an essential contribution to prove and measure the actual suitability of our framework to support DSM.

Overall, the customer segmentation analysis technique proposed in this research is a promising tool that advances traditional methods for inferring residential water and electricity consumer behaviors. The growing deployment of AMI systems worldwide is opening up new opportunities for demand modeling and management in terms of data sampling resolution and data frequency of acquisition. However, the signficantly increased quantity of data will create a burden on existing computational infrastructure without the strategic implementation of more advanced data mining and analysis techniques. We demonstrated the suitability of our method to address such challenges for a subset of roughly 1000 households in California, but future applications should target signficantly larger datasets to test its scalability to broader city, utility-wide, or even regional analysis.

Acknowledgement

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 619172 (SmartH2O: an ICT Platform to leverage on Social Computing for the efficient management of Water Consumption). The authors would like to acknowledge Burbank Water and Power for sharing the anonymized data used in this research.

Appendix

Table A1
List of user psychographic features considered in this study. Source: WaterSmart (2015).

Feature name	Description	Feature nature
Number of house occupants	Number of people usually living in the house	Integer
Home ownership	User owns/rent a house	Binary
Number of toilets	Number of toilets in the house	Integer
Toilet efficiency	Number of toilets using 1.6 gallons per flush, or less	Integer
Grass %	Percentage of grass in the yard	Categorical
Landscape watering 1 (8)	Frequency of landscape watering	Categorical
Landscape watering 2 (8)	User uses drip irrigation	Binary
Landscape watering 3 (8)	User uses overhead sprinklers for irrigation	Binary
Landscape watering 4 (8)	User uses underground irrigation	Binary
Landscape watering 5 (8)	User uses hand-held hoses for irrigation	Binary
Landscape watering 6 (8)	User controls irrigation time with a timer-based controller	Binary
Landscape watering 7 (8)	User controls irrigation time with a weather-based controller	Binary
Landscape watering 8 (8)	User controls irrigation time manually	Binary
Landscape watering schedule 1 (6)	User adjusts the irrigation controller himself/herself	Binary
Landscape watering schedule 2 (6)	Someone else in the household adjusts the irrigation controller	Binary
Landscape watering schedule 3 (6)	The gardener/landscaper adjusts the irrigation controller	Binary
Landscape watering schedule 4 (6)	The homeowner's association adjusts the irrigation controller	Binary
Landscape watering schedule 5 (6)	The property management company adjusts the irrigation controller	Binary
Landscape watering schedule 6 (6)	Frequency of adjustment of the irrigation schedule	Categorical
Water heater type	Type of water heater installed in the house	Categorical
Water heater age	Year of installation of the user's water heater	Categorical
Water-related energy	User estimates the percent of his/her energy bill that goes to water heating	Categorical
Trust in utility 1 (3)	User agrees/disagrees the utility makes it easier for him/her to understand his/her water use	Categorical
Trust in utility 2 (3)	User agrees/disagrees the utility makes it easier for him/her to save money on his/her water bill	Categorical
Trust in utility 3 (3)	User agrees/disagrees the utility makes it easier for him/her take steps to use water efficiently	Categorical
User attitude 1 (5)	User agrees/disagrees he/she usually does things to save water at home	Categorical
User attitude 2 (5)	User agrees/disagrees he/she is interested in spending less on his/her water bill	Categorical
User attitude 3 (5)	User agrees/disagrees saving water helps the environment	Categorical
User attitude 4 (5)	User agrees/disagrees he/she usually talks to others in his/her home about reducing water use	Categorical
User attitude 5 (5)	User agrees/disagrees he/she usually talks with friends about ways to save water	Categorical
Pool/hot tub presence	Home has a pool and/or hot tub	Categorical
Indoor water efficiency actions 1 (7)	User declares he/she already upgraded to a high-efficiency clothes washer, or how likely he/she is going to do so	Categorical
Indoor water efficiency actions 2 (7)	User declares he/she already upgraded to a high-efficiency dishwasher, or how likely he/she is going to do so	Categorical
Indoor water efficiency actions 3 (7)	User declares he/she already upgraded to a more efficient water heater, or how likely he/she is going to do so	Categorical
Indoor water efficiency actions 4 (7)	User declares he/she already replaced an old toilet, or how likely he/she is going to do so	Categorical
Indoor water efficiency actions 5 (7)	User declares he/she already checks his/her showerheads, faucets, and toilets for leaks, or how likely he/she is going to do so	Categorical
Indoor water efficiency actions 6 (7)	User declares he/she already replaced a showerhead with a more efficient model, or how likely he/she is going to do so	Categorical

(*continued on next page*)

Table A1 (*continued*)

Feature name	Description	Feature nature
Indoor water efficiency actions 7 (7)	User declares he/she already replaced a faucet with a more efficient model (or add an aerator), or how likely he/she is going to do so	Categorical
Outdoor water efficiency actions 1 (12)	User declares he/she already adjusts his/her irrigation timer monthly, or how likely he/she is going to do so in the next year	Categorical
Outdoor water efficiency actions 2 (12)	User declares he/she already created an irrigation schedule with short, multiple cycle start times, or how likely he/she is going to do so in the next year	Categorical
Outdoor water efficiency actions 3 (12)	User declares he/she already makes sure water does not run onto sidewalks or driveways, or how likely he/she is going to do so in the next year	Categorical
Outdoor water efficiency actions 4 (12)	User declares he/she already checks for irrigation system leaks, or how likely he/she is going to do so in the next year	Categorical
Outdoor water efficiency actions 5 (12)	User declares he/she already waters at dawn or dusk, or how likely he/she is going to do so in the next year	Categorical
Outdoor water efficiency actions 6 (12)	User declares he/she already replaced sprinkler heads with more efficient nozzles, or how likely he/she is going to do so in the next year	Categorical
Outdoor water efficiency actions 7 (12)	User declares he/she already upgraded to a self-adjusting, weather-based irrigation controller, or how likely he/she is going to do so in the next year	Categorical
Outdoor water efficiency actions 8 (12)	User declares he/she already replaced his/her lawn with drought-tolerant plants, or how likely he/she is going to do so in the next year	Categorical
Outdoor water efficiency actions 9 (12)	User declares he/she already uses drip irrigation for landscaping, or how likely he/she is going to do so in the next year	Categorical
Outdoor water efficiency actions 10 (12)	User declares he/she already installed soil moisture sensors, or how likely he/she is going to do so in the next year	Categorical
Outdoor water efficiency actions 11 (12)	User declares he/she already installed a rain barrel to collect rainwater, or how likely he/she is going to do so in the next year	Categorical
Outdoor water efficiency actions 12 (12)	User declares he/she already installed a home grey water system, or how likely he/she is going to do so in the next year	Categorical

References

Abdallah, A., Rosenberg, D., 2014. Heterogeneous residential water and energy linkages and implications for conservation and management. J. Water Resour. Plan. Manag. 140, 288–297.

Beal, C., Stewart, R.A., Huang, T., Rey, E., 2011. South East Queensland residential end use study. J. Aust. Water Assoc. 38, 80–84.

Beal, C.D., Gurung, T.R., Stewart, R.A., 2016. Demand-side management for supply-side efficiency: modeling tailored strategies for reducing peak residential water demand. Sustain. Prod. Consum. 6, 1–11.

Beal, C.D., Stewart, R.A., Fielding, K., 2013. A novel mixed method smart metering approach to reconciling differences between perceived and actual residential end use water consumption. J. Clean. Prod. 60, 116–128.

Boyle, T., Giurco, D., Mukheibir, P., Liu, A., Moy, C., White, S., Stewart, R., 2013. Intelligent metering for urban water: a review. Water 5, 1052–1081.

Britton, T.C., Stewart, R.A., O'Halloran, K.R., 2013. Smart metering: enabler for rapid and effective post meter leakage identification and water loss management. J. Clean. Prod. 54, 166–176. https://doi.org/10.1016/j.jclepro.2013.05.018.

Cahill, R., Lund, J., 2012. Residential water conservation in Australia and California. J. Water Resour. Plan. Manag. 139, 117–121.

California Department of Water Resources (DWR), 2014. California water Plan Update 2013. http://www.water.ca.gov/waterplan/docs/cwpu2013/Final/Vol2_SouthCoastRR.pdf (Accessed September 15 2017).

Cardell-Oliver, R., 2013. Water use signature patterns for analyzing household consumption using medium resolution meter data. Water Resour. Res. 49, 8589–8599.

Cardell-Oliver, R., Wang, J., Gigney, H., 2016. Smart meter analytics to pinpoint opportunities for reducing household water use. J. Water Resour. Plan. Manag. 142, 04016007.

Carrie Armel, K., Gupta, A., Shrimali, G., Albert, A., 2013. Is disaggregation the holy grail of energy efficiency? The case of electricity. Energy Policy 52, 213–234.

Cominola, A., Giuliani, M., Piga, D., Castelletti, A., Rizzoli, A., Anda, M., 2015a. Modelling residential water consumers' behaviors by feature selection and feature weighting. In: Proceedings of the 36th IAHR World Congress, the Hague (The Netherlands).

Cominola, A., Giuliani, M., Piga, D., Castelletti, A., Rizzoli, A.E., 2015b. Benefits and challenges of using smart meters for advancing residential water demand modeling and management: a review. Environ. Model. Softw. 72, 198–214.

Cominola, A., Moro, L., Riva, A., Giuliani, M., Castelletti, A., 2016. Profiling residential water users' routines by eigenbehavior modelling. In: Proceedings of the 8th International Congress on Environmental Modelling and Software.

Cosgrove, C.E., Cosgrove, W.J., 2012. The Dynamics of Global Water Futures: Driving Forces 2011-2050. Volume 2 of the United Nations World Water Development Report. UNESCO.

Cutter, E., Haley, B., Williams, J., Woo, C., 2014. Cost-effective water-energy nexus: a California case study. Electr. J. 27, 61–68.

DeOreo, W.B., Mayer, P., Dziegielewski, B., Kiefer, J., 2016. Residential End Uses of Water, Version 2: Executive Report. Water Research Foundation.

Desmedt, J., Vekemans, G., Maes, D., 2009. Ensuring effectiveness of information to influence household behaviour. J. Clean. Prod. 17, 455–462.

Eagle, N., Pentland, A.S., 2009. Eigenbehaviors: identifying structure in routine. Behav. Ecol. Sociobiol. 63, 1057–1066.

Escriva-Bou, A., Lund, J.R., Pulido-Velazquez, M., 2015a. Modeling residential water and related energy, carbon footprint and costs in California. Environ. Sci. Policy 50, 270–281.

Escriva-Bou, A., Lund, J.R., Pulido-Velazquez, M., 2015b. Optimal residential water conservation strategies considering related energy in California. Water Resour. Res. 51 (6), 4482–4498.

Espinoza, M., Joye, C., Belmans, R., De Moor, B., 2005. Short-term load forecasting, profile identification, and customer segmentation: a methodology based on periodic time series. IEEE Trans. Power Syst. 20, 1622–1630.

Fischer, C., 2008. Feedback on household electricity consumption: a tool for saving energy? Energy Effic. 1, 79–104.

Friedman, J.H., Fisher, N.I., 1999. Bump hunting in high-dimensional data. Stat. Comput. 9, 123–143.

Gaiser, K., Stroeve, P., 2014. The impact of scheduling appliances and rate structure on bill savings for net-zero energy communities: application to West Village. Appl. Energy 113, 1586–1595. https://doi.org/10.1016/j.apenergy.2013.08.075.

Geller, H., Harrington, P., Rosenfeld, A.H., Tanishima, S., Unander, F., 2006. Polices for increasing energy efficiency: thirty years of experience in OECD countries. Energy policy 34, 556–573.

Gerland, P., Raftery, A.E., ev ikova, H., Li, N., Gu, D., Spoorenberg, T., Alkema, L., Fosdick, B.K., Chunn, J., Lalic, N., Bay, G., Buettner, T., Heilig, G.K., Wilmoth, J., 2014. World population stabilization unlikely this century. Science 346, 234–237. https://doi.org/10.1126/science.1257469.

Gleick, P.H., Haasz, D., Henges-Jeck, C., Srinivas, V., Wolff, G., Cushing, K.K., Mann, A., 2003. Waste Not, Want Not: the Potential for Urban Water Conservation in California. Pacific Institute for Studies in Development, Environment, and Security, Oakland, CA.

Gurung, T.R., Stewart, R.A., Beal, C.D., Sharma, A.K., 2015. Smart meter enabled water end-use demand data: enhanced infrastructure planning of contemporary supply networks platform for the urban water supply networks. J. Clean. Prod. 87, 642–654. https://doi.org/10.1016/j.jclepro.2014.09.054.

Gurung, T.R., Stewart, R.A., Beal, C.D., Sharma, A.K., 2016. Smart meter enabled informatics for economically efficient diversified water supply infrastructure planning. J. Clean. Prod. 135, 1023–1033.

Jolliffe, I., 2002. Principal Component Analysis. Wiley Online Library.

Jorgensen, B.S., Martin, J.F., Pearce, M., Willis, E., 2013. Some difficulties and inconsistencies when using habit strength and reasoned action variables in models of metered household water conservation. J. Environ. Manag. 115, 124–135. https://doi.org/10.1016/j.jenvman.2012.11.008.

Kenway, S.J., Binks, A., Scheidegger, R., Bader, H.P., Pamminger, F., Lant, P., Taimre, T., 2016. Household analysis identifies water-related energy efficiency opportunities. Energy Build. 131, 21–34.

Kenway, S.J., Scheidegger, R., Larsen, T.A., Lant, P., Bader, H.P., 2013. Water-related energy in households: a model designed to understand the current state and

simulate possible measures. Energy Build. 58, 378–389.
Kwac, J., Flora, J., Rajagopal, R., 2014. Household energy consumption segmentation using hourly data. IEEE Trans. Smart Grid 5, 420–430.
Kwakkel, J.H., Jaxa-Rozen, M., 2016. Improving scenario discovery for handling heterogeneous uncertainties and multinomial classified outcomes. Environ. Model. Softw. 79, 311–321.
Kwakkel, J.H., Pruyt, E., 2015. Using system dynamics for grand challenges: the esdma approach. Syst. Res. Behav. Sci. 32, 358–375.
van der Maaten, L., 2014. Accelerating t-sne using tree-based algorithms. J. Mach. Learn. Res. 15, 3221–3245.
van der Maaten, L., Hinton, G., 2008. Visualizing data using t-sne. J. Mach. Learn. Res. 9, 2579–2605.
MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Oakland, CA, USA, pp. 281–297.
Mayer, P.W., DeOreo, W.B., 1999. Residential End Uses of Water. American Water Works Association.
McDonald, R.I., Weber, K., Padowski, J., Flörke, M., Schneider, C., Green, P.A., Gleeson, T., Eckman, S., Lehner, B., Balk, D., Boucher, T., Grill, G., Montgomery, M., 2014. Water on an urban planet: urbanization and the reach of urban water infrastructure. Glob. Environ. Change 27, 96–105. https://doi.org/10.1016/j.gloenvcha.2014.04.022.
Mitchell, D.L., Cubed, M., Chesnutt, T., 2013. Evaluation of East Bay Municipal Utility Districts Pilot of Watersmart Home Water Reports. Policy analysis for the public and private sectors.
Neenan, B., Hemphill, R.C., 2008. Societal benefits of smart metering investments. Electr. J. 21, 32–45. https://doi.org/10.1016/j.tej.2008.09.003.
Nizar, A., Dong, Z., 2009. Identification and detection of electricity customer behaviour irregularities. In: Power Systems Conference and Exposition, 2009. PSCE'09. IEEE/PES, IEEE, pp. 1–10.
Novak, J., Melenhorst, M., Micheel, I., Pasini, C., Fraternali, P., Rizzoli, A.E., 2016. Behaviour change and incentive modelling for water saving: first results from the SmartH2O project. In: Proceedings of the 8th International Congress on Environmental Modelling and Software, Toulouse (France).
Palmgren, C., Stevens, N., Goldberg, M., Barnes, R., Rothkin, K., 2010. 2009 California Residential Appliance Saturation Study, Cec-200-2010-004-es.
Spang, E.S., Loge, F.J., 2015. A high-resolution approach to mapping energy flows through water infrastructure systems. J. Ind. Ecol. 19, 656–665.
Spang, E.S., Loge, F.J., Abernathy, W., Dove, D.R., Tseng, C., Williams, M., et al., 2015. Implementing consumption-based fixed rates in Davis, Calif. J.-Am. Water Works Assoc. 107, E380–E388.
Strbac, G., 2008. Demand side management: benefits and challenges. Energy policy 36, 4419–4426.
UNDESA, 2010. World Urbanization Prospects: the 2014 Revision. Highlights (ST/ESA/SER.A/352). United Nations Department of Economic and Social Affairs Population Division.
US Census Bureau, 2010. Census Interactive Population Search, p. 2010.
Vassileva, I., Campillo, J., 2014. Increasing energy efficiency in low-income households through targeting awareness and behavioral change. Renew. Energy 67, 59–63. https://doi.org/10.1016/j.renene.2013.11.046 renewable Energy for Sustainable Development and Decarbonisation.
Verdú, S.V., Garcia, M.O., Senabre, C., Marín, A.G., Franco, F.J.G., 2006. Classification, filtering, and identification of electrical customer load patterns through the use of self-organizing maps. IEEE Trans. Power Syst. 21, 1672–1682.
WaterSmart, 2015. Participant Entry Survey. WaterSmart Software, San Francisco, CA.
Willis, R.M., Stewart, R.A., Giurco, D.P., Telebpour, M.R., Mousavinejad, A., 2013. End use water consumption in households: impact of socio-demographic factors and efficient devices. J. Clean. Prod. 60, 107–115. https://doi.org/10.1016/j.jclepro.2011.08.006 special Volume: Water, Women, Waste, Wisdom and Wealth.